Class Assignment: MPG tables and charts

Yiru Fei

September 19, 2019

Class assignment:

Using the MPG dataset, which is available with the ggplot2 library, produce a document that includes the following elements: headers, text, tables, and plots.

Tables should include a comparison of city and highway mileage by the class of car and the best three car models for city and highway mileage for all the years in which data is available.

Plot the data displaying as much of the data as you can. Put continuous variables on the axes and include a locally smoothed regression line to show the relationship with mileage. Also make a box plot comparing city and highway MPG by class of car.

Tables

This example is flawed even though the code works. Explain. And now that you know how to produce the table can you improve the code by using an **apply** functional?

```
data(mpg)
Class <- c("2seater","compact","midsize","minivan","pickup","subcompact","suv")
cty_mean <- by(mpg, mpg$class, function(x) c(mean(x$cty)))
hwy_mean <- by(mpg, mpg$class, function(x) c(mean(x$hwy)))
cty_mean <- round(cty_mean, 2)                    ## round the results to 2 decimal pla
hwy_mean <- round(hwy_mean, 2)


tbl_1 <- cbind( Class, cty_mean,hwy_mean)      ## form an array for the table
rownames(tbl_1) <- NULL


kable(tbl_1, format = "latex", booktabs=TRUE, digits = 2,     ## call kable to make the table
      col.names = c( "Class","City", "Highway"),
      caption = "Mean City and Highway MPG by Car Class" )
```

The next table is a bit tricky.

```
##  You probably want to get started with a little investigation.
##  Using unique(mpg$year), you find out that there are only two years - 1999, 2008.


mpg_99=mpg[which(mpg$year==1999),]
mpg_08=mpg[which(mpg$year==2008),]

c_1 = c("new beetle","civic","corolla")
c_2 = c("new beetle","corolla","civic")
c_3 = c("corolla","civic","gti")
c_4 = c("corolla","civic","camry")
```

Table 1: Mean City and Highway MPG by Car Class

Class	City	Highway
2seater	15.4	24.8
compact	20.13	28.3
midsize	18.76	27.29
minivan	15.82	22.36
pickup	13	16.88
subcompact	20.37	28.14
suv	13.5	18.13

```r
mod_99_cty_mean <- by(mpg_99, mpg_99$model, function(x) c(mean(x$cty)))
mod_08_cty_mean <- by(mpg_08, mpg_08$model, function(x) c(mean(x$cty)))

mod_99_hwy_mean <- by(mpg_99, mpg_99$model, function(x) c(mean(x$hwy)))
mod_08_hwy_mean <- by(mpg_08, mpg_08$model, function(x) c(mean(x$hwy)))

ind_srt_99_cty <- order(mod_99_cty_mean)
miles_cty_99 <- rev(mod_99_cty_mean[ind_srt_99_cty])



ind_srt_08_cty <- order(mod_08_cty_mean)
miles_cty_08 <- rev(mod_08_cty_mean[ind_srt_08_cty])



ind_srt_99_hwy <- order(mod_99_hwy_mean)
miles_hwy_99 <- rev(mod_99_hwy_mean[ind_srt_99_hwy])



ind_srt_08_hwy <- order(mod_08_hwy_mean)
miles_hwy_08 <- rev(mod_08_hwy_mean[ind_srt_08_hwy])



miles_cty_08 <- round(miles_cty_08, 2)
miles_cty_99 <- round(miles_cty_99, 2)

miles_hwy_08 <- round(miles_hwy_08, 2)
miles_hwy_99 <- round(miles_hwy_99, 2)



tbl_2 <- cbind(c_1,miles_cty_99[1:3],

               c_2,miles_hwy_99[1:3],

               c_3,miles_cty_08[1:3],

               c_4,miles_hwy_08[1:3])
```

Table 2: Top 3 MPG Performing Cars: 1999, 2008

City 1999		Highway 1999		City 2008		Highway 2008	
	Model	Milage	Model	Milage	Model	Milage	Model	Milage
new beetle	new beetle	26	new beetle	35	corolla	27	corolla	36
civic	civic	24.8	corolla	32.67	civic	24	civic	33.75
corolla	corolla	24.67	civic	31.6	gti	21.5	camry	30

```r
colnames(tbl_2) <- c('Model', 'Milage',
                     "Model", "Milage",
                     'Model', 'Milage',
                     "Model", "Milage"
                     )

kable(tbl_2, digits = 2, format = "latex", booktabs=TRUE, ,caption = "Top 3 MPG Performing Cars: 1999, 
  add_header_above(c("City 1999"=2,
                     "Highway 1999"=2,
                     "City 2008"=2,
                   "Highway 2008"=2))

library(ggplot2)

ggplot(mpg) +
 aes(x = displ, y = cty) +
 geom_point(aes(color=class),size=2) + geom_smooth() +
 scale_color_hue() +
 theme_minimal(base_size = 16) +
 facet_wrap(vars(year)) +
  labs(x = "Engine Displacement",y = "City Mileage" ,title = "City MPG by Class of Car: 1999, 2008")
```
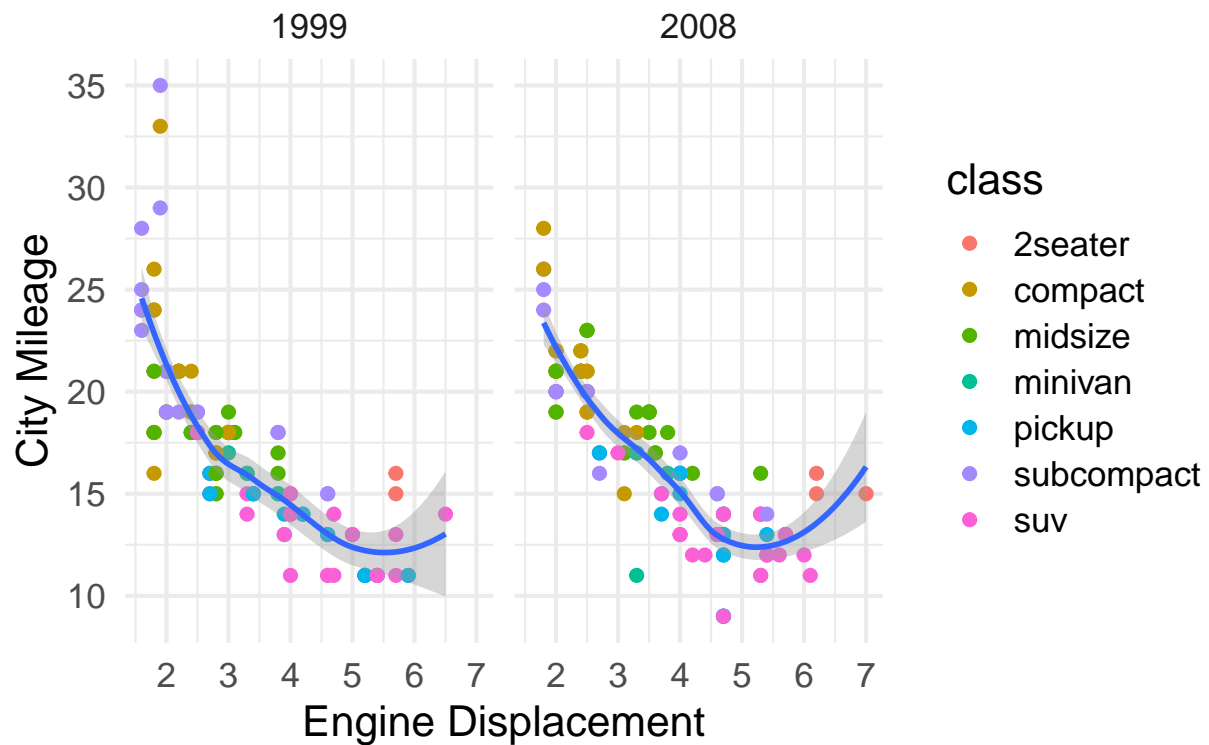
```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```
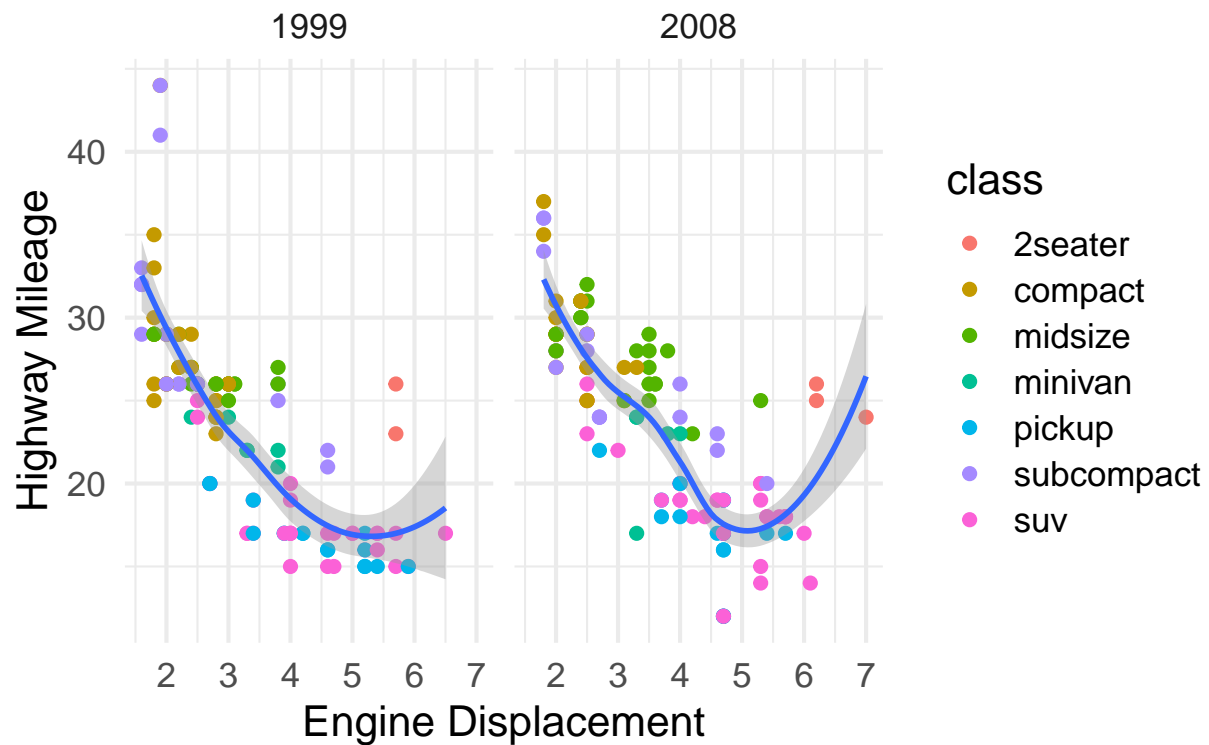
City MPG by Class of Car: 1999, 2008


```
ggplot(mpg) +
aes(x = displ, y = hwy) +
geom_point(aes(color=class),size=2) + geom_smooth() +
scale_color_hue() +
theme_minimal(base_size = 16) +
facet_wrap(vars(year)) +
 labs(x = "Engine Displacement", y = "Highway Mileage", title = "Highway MPG by Class of Car: 1999, 200
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

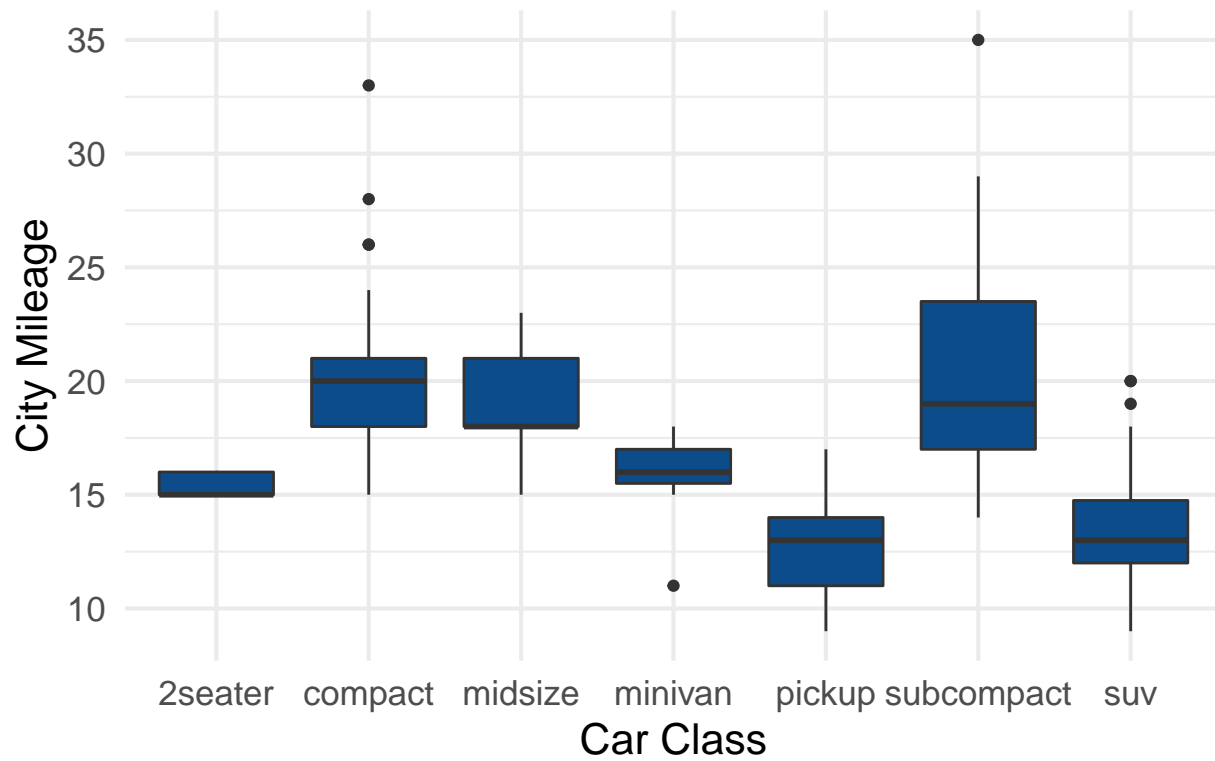# Highway MPG by Class of Car: 1999, 2008


```
##esquisser(mpg)
```

Boxplots

```
ggplot(mpg) +
 aes(x = class, y = cty) +
 geom_boxplot(fill = "#0c4c8a") +
 theme_minimal(base_size=16) +
  labs(x = "Car Class", y = "City Mileage", title = "City MPG by Class of Car: 1999, 2008")
```

City MPG by Class of Car: 1999, 2008


```
ggplot(mpg) +
 aes(x = class, y = hwy) +
 geom_boxplot(fill = "#0c4c8a") +
 theme_minimal(base_size=16) +
  labs(x = "Car Class", y = "Highway Mileage", title = "Highway MPG by Class of Car: 1999, 2008")
```

Highway MPG by Class of Car: 1999, 2008